

## SPECIFICATION

## TITLE OF THE INVENTION

APPARATUS WITH DUAL WRITING FUNCTION, AND STORAGE CONTROL .

## APPARATUS

## 5 BACKGROUND OF THE INVENTION

## 1) Field of the Invention

The present invention relates to an apparatus having a dual writing function (multi-writing function) to write data, transferred from an external apparatus such as a host, in two or more cache memories, and more particularly to a storage control unit provided between a physical device (for example, a magnetic disk unit) and a host for controlling the access to the physical device by the host. The present invention is suitable for use in RAID (Redundant Arrays of Inexpensive Disk).

## 2) Description of the Related Art

In general, a storage apparatus serving as an external storage unit is connected through a network or the like to a host (computer). This storage apparatus is for, in accordance with the access (input/output request, I/O request) from the host, writing data from the host in a physical device (for example, a disk unit) or reading out data requested by the host from the physical device to transfer it to the host, and it is made up of a disk unit and a storage control unit located between the disk unit and the host for controlling the access from the host to the disk unit.

Usually, this storage control unit includes at least a host interface module for controlling the interface to the host, a disk interface module for controlling the interface to the disk unit, and two or more management modules  
5 for generally managing the entire storage control unit.

In addition, a cache memory is mounted on the management modules to temporarily store data to be written (which sometimes will be referred to equally as to written data) from the host into the disk unit or data to be read out  
10 from the disk unit into the host, and each of the management modules is also made to manage this cache memory.

As an interface bus for a connection between the host interface module and the management modules or an interface bus for a connection between the disk interface module and  
15 the management modules, there is used a standardized bus, for example, PCI (Peripheral Component Interconnect) bus.

In the storage control unit thus configured, in reading out data from the disk unit into the host, the data to be read is first transferred from a disk unit preserving the  
20 data to the disk interface module and then shifted from the disk interface module through the PCI bus to the cache memory of the management module to be temporarily stored therein. Following this, the data to be read, existing on the cache memory of the management module, is transferred  
25 through the PCI bus to the host interface module and then read out from the host interface module into the host.

Conversely, in writing data from the host into the

disk unit, the data to be written is transferred from the host to the host interface module and then shifted from the host interface module through the PCI bus to the cache memory of the management module to be temporarily stored  
5 therein. At this time, the data to be written, which has a high frequency of the access from the host, is retained in the cache memory, and if the frequency becomes low, the data is transferred from the cache memory through the PCI bus to the disk interface module to be written in a  
10 predetermined disk unit by the disk interface module.

Moreover, in general, in writing data from the host into the storage apparatus, for the purpose of preventing user's data before the writing in a disk unit such as a magnetic disk unit from disappearing due to troubles of  
15 a hardware (management module) including the cache memory, or the like, in the interior of the storage apparatus, the user's data transferred from the host is doubly written in two cache memories (the same data is written in a plurality of hardwares) (for example, see Japanese patent Laid-Open  
20 Nos. HEI 07-160432, 05-189314 and 07-20994).

In the storage control unit arranged as mentioned above, in a case in which data is dual-written in cache memories of two management modules (for example, CM-A and CM-B), the host interface module, when receiving data to be written  
25 from the host, usually transfers this data through the PCI bus to a designated address of the CM-A (cache memory A) to write it thereat and further transfers the same data

through the PCI bus to a designated address of the CM-B (cache memory B) to write it thereat. That is, the same data is transferred twice between the host interface module and the two management modules.

5           However, for the dual-writing of data, the conventional storage control unit requires that the same data is transferred twice from one host interface module to two management modules through the use of a standardized bus having a low transfer rate (speed), and this is not  
10       preferable in processing performance. Therefore, a desired solution to this problem involves carrying out the dual-writing of data in cache memories at a higher rate by one address designation for the enhancement of the processing performance.

15           In addition, in a case in which a mechanism for the dual-writing is simply incorporated into a hardware, there is a need to make the management so that data is written at the identical address on cache memories of two management modules (CM-A, CM-B). However, this dual-writing causes  
20       a severe limitation on the data allocation (data layout) in the cache memories to lower the degree of freedom of the data allocation and, hence, useless areas develop in the cache memories. For this reason, there is also a desire to increase the degree of freedom of the data allocation  
25       (cache management) in the cache memories for efficient use of the cache memories.

SUMMARY OF THE INVENTION

The present invention has been developed with a view to eliminating these problems, and it is therefore an object of the invention to carry out the dual-writing of data in cache memories at a higher rate by a single address designation for the enhancement of the processing performance and further to increase the degree of freedom of the cache management for realizing the efficient use of the cache memories.

For this purpose, in accordance with an aspect of the present invention, there is provided an apparatus with a dual-writing function comprising a first module for controlling an interface to an external apparatus, a plurality of second modules each having a cache memory and a bridge module connected through an interface bus to the first and second modules for accomplishing a connection between the first module and the second module for data transfer therebetween, the first module including address designation means for producing addressing information to designate two written-in destinations for writing data to be written, which is received from the external apparatus, through the bridge module into the cache memories of two of the plurality of second modules, and the bridge module including address production means for analyzing the addressing information, which is received together with the data to be written from the first module, to produce two transferred-to addresses for the designation of the two second modules having the cache memories in which the

data is to be actually written and to produce written-in addresses in the cache memories and further including data transfer control means for controlling data transfer from the bridge module to the second modules so that, after the  
5 data to be written is transferred to the two second modules corresponding to the two transferred-to addresses, the data to be written is written at the written-in address in the cache memory of each of the two second modules.

In this apparatus, it is also appropriate that each  
10 of the second modules includes management means for managing information on the second module which is in mirror relation to this second module, and for managing the association between a master area address in the cache memory of this second module and a mirror area address in the cache memory  
15 of the second module being in mirror relation to this second module, and the address designation means of the first module produces the addressing information on the basis of information acquired from the management means of one of the two second modules.

20 In addition, in accordance with another aspect of the present invention, there is provided a storage control apparatus placed between a disk unit and a host for controlling access to the disk unit by the host, and comprising a disk interface module for controlling an  
25 interface to the disk unit, a host interface module for controlling an interface to the host, a plurality of management modules, each including a cache memory, for

generally controlling the entire apparatus, and a bridge module connected through interface buses to the disk interface module, the host interface module and the management modules for making connections among the disk interface module, the host interface module and the management modules for data transfer among the modules, the host interface module including address designation means for producing addressing information to designate two written-in destinations for writing data to be written, which is received from the host, through the bridge module into the cache memories of two of the plurality of management modules, and the bridge module including address production means for analyzing the addressing information, which is received together with the data to be written from the host interface module, to produce two transferred-to addresses for the designation of the two management modules having the cache memories in which the data is to be actually written and to produce written-in addresses in the cache memories and further including data transfer control means for controlling data transfer from the bridge module to the management modules so that, after the data to be written is transferred to the two management modules corresponding to the two transferred-to addresses, the data to be written is written at the written-in address in the cache memory of each of the two management modules.

In this arrangement, it is also appropriate that the address designation means designates, in the addressing

information, a page address in the cache memory of each of the management modules and an offset address in a page designated by the page address, as the written-in address for the data to be written in the cache memory. Moreover, 5 it is also appropriate that each of the management modules includes management means for managing information on the management module which is in mirror relation to this management module and for managing the association between a master area address in the cache memory in this management 10 module and a mirror area address in the cache memory of the management module being in the mirror relation to this management module, and the address designation means of the host interface module produces the addressing information on the basis of information acquired from the 15 management means of one of the two management modules.

With the aforesaid apparatus with a dual-writing function and storage control apparatus according to the present invention, the address designation means of the first module (host interface module) produces the addressing 20 information for the designation of two written-in destinations, and in the bridge module, the address production means produces, on the basis of the addressing information, two transferred-to addresses (two second modules (management modules)) and the written-in addresses 25 in the respective cache memories. Moreover, the data transfer control means of the bridge module transfers the data to be written to the two second modules (management



modules) corresponding to the two transferred-to addresses, with it being written at the written-in address of the cache memory of each of the second modules (management modules).

Therefore, when the first module (host interface  
5 module) once carries out the address designation to make the data transfer from the first module (host interface module) to the bridge module, the data to be written is transferred through the bridge module, which provides a higher transfer rate than that of a standardized bus, to  
10 the cache memories of the two second modules (management modules) to be doubly written therein, which enables the dual-writing in the cache memories to be done at a high rate for considerable improvement of the processing performance.

15 Moreover, since the address designation means of the first module (host interface module) designates, in the addressing information, a page address in the cache memory of each of the management modules and an offset address in each page as the written-in address for the data to be  
20 written in each of the cache memories, the data to be written can be transferred to separate addresses in the cache memories of the two second modules (management modules) to be written therein, thereby increasing the degree of freedom of the cache management and achieving the efficient  
25 use of the cache memories.

Still moreover, since each of the second modules (management modules) is equipped with the management means

and this management means manages the information on a second module (management module) which is in mirror relation to this second module and manages the association between a master area address in the cache memory of this second module (management module) and a mirror area address in the cache memory of the management module being in mirror relation to this second module (management module), so the address designation means of the first module (host interface module) can produce the addressing information on the basis of the information acquired from the management means in one of the two second modules (management modules) without making communications between the two second modules (management modules) which are in mirror relation to each other.

Yet moreover, since the management means manages the mirror area of the second module (management module) in mirror relation, one second module (management module) can efficiently utilize the mirror area of the other management module without making communications between the two second modules (management modules) which are in mirror relation to each other. For example, in a case in which the capacity of the master area in the cache memory of one second module (management module) runs short, on the basis of the situation of the management by the aforesaid management means, the utilization of the mirror area in the cache memory of the other second module (management module) is feasible, which enables the efficient utilization of both the master area

and mirror area in the cache memory.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram showing a minimum basic configuration of a storage apparatus (storage control unit) according to an embodiment of the present invention;

FIG. 2 is an illustration of a format for addressing information to be produced in the storage control unit (host interface module) shown in FIG. 1;

FIG. 3 is an illustration of an arrangement of a cache memory according to this embodiment;

FIG. 4 is an illustration of the association between a master area (local area) and a mirror area in the cache memory according to this embodiment;

FIG. 5 is an illustration useful for explaining one example of a management operation for the cache memory according to this embodiment;

FIG. 6 is a block diagram showing an example of a concrete configuration of a storage apparatus (storage control unit) according to an embodiment of the present invention;

FIG. 7 is a block diagram showing a concrete example of an essential part of a bridge module of the storage control unit shown in FIG. 6; and

FIG. 8 is an illustration of, in the storage control unit shown in FIGS. 6 and 7, a format for addressing information to be produced in a host interface module and a format for address information to be produced from the

same addressing information in a PCI bridge module.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

Embodiments of the present invention will be described hereinbelow with reference to the drawings.

5           FIG. 1 is a block diagram showing a minimum basic configuration of a storage apparatus (storage control unit) according to an embodiment of the present invention. In FIG. 1, the storage apparatus, generally denoted at reference numeral 1, is designed to, in accordance with  
10           the access (input/output request, I/O request) from a host 4, write data from the host 4 or read out data requested by the host 4 for transferring it to the host 4. This storage apparatus 1 is made up of a disk unit (physical device) 2 and a storage control unit (storage control apparatus) 3 provided between the disk unit 2 and the host 4 for  
15           controlling the access to the disk unit 2 by the host 4.

          The storage control unit 3 is composed of a disk interface module 10, a host interface module 20, management modules 30 and a PCI bridge module 40. In the storage control  
20           unit 3 having a minimum basic configuration shown in FIG. 1, each of the disk interface module 10, the host interface module 20 and the PCI bridge module 40 is one in number, while the management modules 30 are two in number.

          The disk interface module 10 is for controlling the  
25           interface (data transfer) through a disk interface bus 54 to the disk unit 2.

          The host interface module (first module) 20 is for

controlling the interface (data transfer) through a fiber channel interface bus 50 to the host (external apparatus)

4. This host interface module 20 has a function as an address designation means 21 which will be described later.

5        Each of the management modules (second modules) 30 is for generally controlling the entire storage control unit 3, and it is equipped with a cache memory 31 for temporarily storing data from the host 4 to be written in the disk unit 2 or data to be read out from the disk unit  
10       2 into the host 4, and the management module 30 also manages this cache memory 31. The management module 30 has a function as a management means 32 which will be mentioned later.

      The PCI bridge module (bridge module) 40 is connected  
15       through PCI buses (interface buses) 51, 52 and 53 serving as standardized buses to the disk interface module 10, the host interface module 20 and the management modules 30 for achieving the data transfer among these disk interface module 10, the host interface module 20 and management  
20       modules 30. This PCI bridge module 40 has functions as an address production means 41 and a data transfer control means 42, which will be described later. Incidentally, although a general PCI-PCI bridge is designed to make the one-to-one connection, the PCI bridge module 40 to be used  
25       in this configuration is designed to make n-to-n (plurality: plurality) connections.

      The duplication of written data is accomplished by

two management modules 30 which are in mirror relation to each other. In the storage control unit 1 with a minimum basic configuration shown in FIG. 1, the management modules 30 are two in number, and these two management modules 30 are set in mirror relation to each other, and one management module 30 is handled as a master CM while the other management module 30 is handled as a mirror CM (pair CM).

Referring to FIGs. 3 to 5, a description will be given hereinbelow of the cache memory 31 in each of the management modules 30. FIG. 3 is an illustration of an arrangement of the cache memory 31 according to this embodiment, FIG. 4 is an illustration of the association of a master area (local area) and a mirror area in the cache memory 31 according to this embodiment, and FIG. 5 is an illustration useful for explaining one example of a management operation of the cache memory 31 according to this embodiment.

As FIG. 3 shows, an area (cache area) in the cache memory 31 is partitioned into cache pages each having a predetermined size (for example, 2 MB (Mega Byte)), with an area (corresponding to two pages) corresponding to 4 MB being managed as one unit. Moreover, for the arrangement/management of each page, 31 CBEs (Cache Bundle control Element) each having  $520 \text{ bytes} \times 128 \text{ blocks} = 66560 \text{ bytes}$  are allocated thereto. The management of each cache page is made by connecting a non-used CBE to a free link group. The CBE in use is managed by connecting it to an LRU (Least Recently Used) link and a hash entry. However, the

CBEs in the mirror area are not connected to a free link.

In addition, as shown in FIGs. 3 and 4, in the cache area, the master areas (master CBE group) and the mirror areas (mirror CBE group) are alternately disposed in units  
5 of 4 MB to perform the management as a two-division memory format. In FIG. 4, the association between the master areas (local areas) and the mirror areas in the cache memories 31 of the two management modules 30 (CM-A and CM-B) which are in mirror relation to each other is indicated by arrows.

10 Still additionally, in this embodiment, the master area (local area) in the cache memory 31 of each of the management modules 30 is used to transfer and keep its own data for I/O requests accepted by the management module 30, and the mirror area in the cache memory 31 of each of  
15 the management modules 30 is basically used to transfer and keep duplicated data for the other management module 30. Yet additionally, in this embodiment, the mirror area in the cache memory 31 of each of the management modules 30 is also used to transfer readout data for I/O requests  
20 accepted by the other management module 30 as mentioned later with reference to FIG. 5. That is, in this embodiment, the mirror area of the cache memory 31 of each of the management modules 30 is managed by the other management module 30 which is positioned backwardly, and this  
25 management module 30 cannot use it directly. For example, the master area of CM-A is used as temporal storage to read or write data for I/O requests accepted by CM-A, and the

mirror area of CM-A is used as temporal storage to write duplicated data for write requests accepted by CM-B and to read data for read requests accepted by CM-B. The all of mirror area of CM-A is managed by CM-B.

5           In this connection, in the master area/mirror area in each cache memory 31, the aforesaid readout data transfer memory area is established and, hence, a minimum capacity to be used as only that memory area is set in advance. This minimum capacity is defined/set as a "Read cache capacity"  
10 of a tuning parameter. For example, the default value of this minimum capacity is 64 MB, and if it is set at 64 MB, a readout data transfer memory area of approximately 1000 pages (exactly, 992 pages) is securable in the form of CBEs.

          The master area and the mirror area in each of the  
15 cache memories 31 are restricted by the physical partitioning. This physical restriction enables the mirror cache (the mirror area of the mirror CM side cache memory 31) to be managed by the master CM side management means 32 (which will be mentioned later). At this time,  
20 the boundary between the master area and the mirror area on the physical arrangement is set in consideration of the cache page of the hardware. This allows the implementation of the dual-writing (forking write) according to the present invention. In the forking write according to this  
25 embodiment, as will be mentioned later, although the cache page address and the offset value (offset address) at that address are designated by the address designation means



21 of the host interface module 20, the offset values in the master area and the mirror area are required to be set at the same value due to the aforesaid hardware physical restriction.

5           In the storage control unit 3 according to this embodiment, for managing the above-mentioned cache memories 31, each of the management modules 30 has a function as a management means (cache management firmware) 32.

          This management means 32 manages the information  
10   (number specifying a PCI bus of the mirror CM, or the like) on the management module (mirror CM) which is in mirror relation to the management module 30 (master CM) including this management means 32, and manages the association (this association is determined in advance) between the master  
15   area address in the master CM side cache memory 31 and the mirror area address in the mirror CM side cache memory 31 so that the mirror cache (the mirror area of the mirror CM side cache memory 31) is manageable from the master CM side. That is, the use/non-use of the CBEs in the mirror  
20   cache is manageable by the master CM side management means 32.

          Furthermore, in this embodiment, since each of the management modules 30 is made such that the mirror cache is manageable from the master CM side through the use of  
25   the aforesaid management means 32, as shown in FIG. 5, in a case in which one management module 30 (CM-A) temporarily preserves, in the cache memory 31, the data read out from

the disk unit 2 through the disk interface module 10 and the PCI bridge module 40 (omitted in FIG. 5) in accordance with a readout request (I/O request) by the host 4, even if the capacity of the master area of this cache memory 5 31 runs short, as indicated by an arrow A1 (staging processing) in FIG. 5, on the basis of the management situation in the management means 32, the management module 30 (CM-A) can preserve the aforesaid readout data in the mirror area (the aforesaid readout data transfer memory area secured in advance) in the cache memory 31 of the 10 management module 30 (CM-B) which is in mirror relation to the management module 30 (CM-A). As indicated by an arrow A2 (host readout processing) in FIG. 5, the readout data preserved in the mirror cache in this way is transferred 15 from the mirror area of the CM-B through the PCI bridge module 40 and the host interface module 20 to the host 4.

The CBE (the corresponding CBE in the mirror area of the management module 30 being in mirror relation to that management module 30) paired with the CBE in use for the 20 readout data in the master area of the cache memory 31 of the management module 30 is not used for the duplication of the written data, and it is left as a free area. If such an area is put in practical use, also for the I/O processing on a readout request, the cache memory 31 becomes usable 25 with a using efficiency of approximately 100%. In this embodiment, the aforesaid free area in the mirror area of the mirror CM can be used as a readout data transfer memory

area through the use of the function of the management means 32 in each of the management modules 30.

However, for realizing this function, between the management modules 30 and the disk interface module 10 or the host interface module 20, there is provided an interface accessible not only to the cache memory of an in-charge management module 30 (CM-A in the example shown in FIG. 5) set in advance for these interface modules 10 and 20 but also to the cache memory 31 of the management module 30 (CM-B in the example shown in FIG. 5) different from this in-charge management module 30. Through this interface, an I/O request cache address and a PCI number to a designated address of the management module 30 are designated from the management module 30 side. Moreover, when the interface modules 10 and 20 see the cache memory 31, the PCI number is additionally set up.

The storage control unit 3 according to this embodiment is equipped with the aforesaid management modules 30 (cache memory 31 and management means 32) to carry out the written data dual-writing (duplication) referred to as forking writing. Secondly, a description will be given hereinbelow of a function (function of the address designation means 21 of the host interface module 20, functions of the address production means 41 and data transfer control means 42 of the PCI bridge module 40) for carrying out the forking writing operation.

The address designation means 21 of the host interface

module 20 produces, on the basis of information acquired from the management means 32 in one (in this case, CM-A) of the two management modules 30, addressing information for designating two written-in destinations in which the data to be written from the host 4 is dual-written in the cache memories 32 of the two management modules 30 through the PCI bridge module 40.

FIG. 2 is an illustration of a format for the addressing information to be produced by the address designation means 21. As shown in FIG. 2, in the aforesaid addressing information, the following (1) to (4) are designated as addresses (PCI addresses) to be instructed to the PCI bus 52. That is,

- (1) Base Address (BA in FIG. 2);
- (2) written-in page address (PA-A; primary page in FIG. 2) in the cache memory 31 of CM-A;
- (3) written-in page address (PA-B; secondary page in FIG. 2) in the cache memory 31 of CM-B; and
- (4) offset address (offset value; offset address in FIG. 2) indicative of a written head position in each cache page designated by the written-in page address of the aforesaid (2) and (3).

Incidentally, in the storage control unit 3 with a minimum basic configuration shown in FIG. 1, since the management modules 30 provided are only two in number and these two management modules 30 are always in the mirror relation to each other, unlike an example described later

with reference to FIGs. 6 to 8, it is not necessary that the information (two transferred-to addresses) for specifying the two transferred-to management modules 30 to which the data to be written is transferred be designated  
5 in the aforesaid addressing information, and it may be fixedly retained in the interior of the PCI bridge module 40.

Moreover, the base address (BA) of the aforesaid (1) is information for making an instruction to the PCI bridge  
10 module 40 for, in the present I/O request (writing request), the data to be written being transferred from the host interface module 20 through the PCI bridge module 40 to the management module 30. In more concretely, for example, an instruction to the PCI bridge module 40 is made such  
15 that, when the base address is "01", the data to be written is transferred to only the CM-A to be written therein, and when the base address is "10", the data to be written is transferred to only the CM-B to be written therein, and when the base address is "11", the data to be written is  
20 transferred to both the CM-A and CM-B to be written therein. Accordingly, in a case in which "11" is designated as the base address, the forking writing (dual-writing) according to this embodiment is implemented through the use of the functions of the address production means 41 and the data  
25 transfer control means 42 in the PCI bridge module 40.

The address production means 41 of the PCI bridge module 40 analyzes the addressing information received together

with the data to be written from the host interface module 20, and produces two transferred-to addresses for designating the two management modules (CM-A, CM-B) having the cache memories 31 in which the written data is actually written and written-in addresses in the respective cache memories 31.

As mentioned above, in the storage control unit 3 with a minimum basic configuration shown in FIG. 1, since the aforesaid two transferred-to addresses are not designated in the aforesaid addressing information but being fixedly retained in the interior of the PCI bridge module 40, it is produced by reading out the information retained therein. Moreover, the following addresses (a1) and (a2) are produced as the written-in address on the cache memory of the CM-A and the written-in address on the cache memory 31 of the CM-B, respectively.

(a1) (PA-A of CM-A) + (offset value)

(a2) (PA-B of CM-B) + (offset value)

Thus, the data to be written is stored at different addresses in the cache memories 31 of the master CM and the mirror CM, and the offsets (written positions) in the cache pages are managed to be equal to each other irrespective of the difference in page address.

The data transfer control means 42 is for controlling the data transfer from the PCI bridge module 40 to the management modules 30 to transfer the same data to be written to the two management modules (CM-A, CM-B) corresponding

to the two transferred-to addresses so that it is written at the aforesaid written-in addresses (a1) and (a2), produced by the address production means 41, in the cache memories 31 of the management modules 30.

5           With the above-mentioned arrangement, in the storage control unit 3 with the minimum basic configuration shown in FIG. 1, when the host interface module 20 receives a data writing request as the I/O request from the host 4 through fiber channel interface bus 50 (see an arrow A11  
10 in FIG. 1), the address designation means 21 produces the addressing information (see FIG. 2) for designating two written-in destinations. This addressing information, together with the data to be written, is transferred from the host interface module 20 through the PCI bus 52 to the  
15 PCI bridge module 40 (see an arrow A12 in FIG. 1).

          In the PCI bridge module 40, when recognizing a written-in command with the addressing information, the address production means 41 produces, on the basis of the addressing information from the host interface module 20,  
20 two transferred-to addresses (addresses designating CM-A and CM-B) and written-in addresses (a1) and (a2) in the cache memories 31. Moreover, the data transfer control means 42 transfers the same data to be written to the two management modules 30 corresponding to the aforesaid two  
25 transferred-to addresses, with it being written at the written-in addresses (a1) and (a2) in the cache memories 31 of the management modules 30 (see arrows A13 and A14

in FIG. 1).

In this way, the forking writing (dual-writing of the data to be written) is implemented as indicated by the arrows A11 to A14 in FIG. 1, and the same written data is written in a CBE of a cache page of the master area of the cache memory 31 of the CM-A and in a CBE (paired CBE) of a cache page of the mirror area of the cache memory 31 of the CM-B which is associated with the aforesaid master area cache page CBE in advance.

Furthermore, referring to FIGs. 6 to 8, a description will be given hereinbelow of a concrete example of a configuration of the storage apparatus (storage control unit) according to an embodiment of the present invention. FIG. 6 is a block diagram showing a concrete example of a configuration of the storage apparatus 1 (storage control unit 3), FIG. 7 is a block diagram showing a concrete example of an essential part of the PCI bridge module 40 in the storage control unit 3 shown in FIG. 6, and FIG. 8 is an illustration of a format for addressing information to be produced by the host interface module 20 (address designation means 21) in the storage control unit 3 shown in FIG. 7 and a format for address information to be produced from this addressing information by the PCI bridge module 40 (address production means 41). In FIGs. 6 and 7, the same reference numerals as those used above represent the same or approximately same parts, and the description thereof will be omitted for brevity.



As FIG. 6 shows, the storage control unit 3 of the storage apparatus 1 is made up of eight host interface modules 20, four management modules 30, four disk interface modules 10 and two PCI bridge modules 40 and, for example, eight disk units 2 are connected through disk interface buses 54 to the four disk interface modules 10.

In the storage control unit 3 shown in FIG. 6, each of the host interface modules 20 also has a function as the aforesaid address designation means 21 (omitted in FIG. 6). This address designation means 21 produces, on the basis of the information acquired from the management means 32 of one (CM-A, master CM) of the two management modules 30, the addressing information to designate two written-in destinations so that the data to be written the host interface module 20 receives from the host 4 is doubly written through the PCI bridge module 40 in the cache memories 31 of the two management modules 30. However, in the host interface module 20 (address designation means 21) shown in FIG. 6, as indicated by (A) in FIG. 8, in the aforesaid addressing information, the following addresses (1) to (7) are designated as the addresses (PCI addresses) to be instructed to the PCI buses 52. That is,

- (1) base address (BA in FIG. 8);
- (2) written-in page address (PA-A; primary page in FIG. 8) in the cache memory 31 of CM-A;
- (3) written-in page address (PA-B; secondary page in FIG. 8) in the cache memory 31 of CM-B;

(4) offset address indicative of a written head position in each cache page designated by the aforesaid written-in page addresses of (2) and (3);

(5) number for specifying a PCI bus of CM-A (specific  
5 information for specifying the master CM having the cache memory 31 in which data is to be actually written; P-PCI in FIG. 8);

(6) number for specifying a PCI bus of CM-B (specific  
information for specifying the mirror CM having the cache  
10 memory 31 in which data is to be actually written; S-PCI in FIG. 8); and

(7) arbitrary number given for specifying an operation (TID (Transaction ID)).

Since the storage control unit 3 shown in FIG. 6 uses  
15 four management modules 30, there is a need to designate, of these four management modules 30, two management modules 30 which is in mirror relation to each other. Therefore, the address designation means 21 is made to designate the aforesaid information of (5) and (6).

20 Moreover, the aforesaid base address (BA) of (1) fulfills a function similar to that shown in FIG. 2, and this BA is used for an instruction to the PCI bridge module 40. For example, when the base address is "10", the data to be written is transferred to only the management module  
25 30 designated in the CM-A field (the aforesaid (5)) to be written in only the CM-A with no regard to the designation value in the CM-B field (the aforesaid (6)), and when the

base address is "11", the data to be written is transferred to both the CM-A and CM-B to be written therein. Therefore, in this transfer instruction (dual-writing instruction) method, when "11" is designated as the base address, the  
5 forking writing (dual-writing) is implemented through the use of the functions of the address production means 41 and the data transfer control means 42 in the PCI bridge module 40.

In addition to the transfer instruction (dual-writing  
10 instruction) method using BA mentioned above, it is also possible to employ a method in which, without using BA, the transfer and writing are basically always made with respect to two management modules 30 (CM-A and CM-B) designated in the aforesaid fields of (5) and (6) and, only  
15 when the CM-A designation value and the CM-B designation value are equal to each other, the transfer and writing is made with respect to one management module 30.

In the storage control unit 3 shown in FIG. 6, as shown in FIG. 7, the PCI bridge module 40 is equipped with a bus  
20 control device 43, an address analysis/conversion device 44, internal transfer control devices 45, buffers 46 and 47, transfer instruction devices 48 and bus control devices 49.

The bus control devices 43 and 49 are for controlling  
25 the PCI buses 52 and 53 (interface control).

The address analysis/conversion device 44 fulfills a function similar to that of the aforesaid address

production means 41, and analyzes the addressing information, shown in (A) of FIG. 8, received together with the data to be written from the host interface module 20 to produce two transferred-to addresses for designating two management modules 30 (CM-A and CM-B) having the cache memories 31 in which the data is actually written and written-in addresses in the respective cache memories 31.

The two transferred-to addresses are produced by reading out two PCI bus numbers (P-PCI and S-PCI) in the aforesaid (5) and (6) fields from the addressing information shown in (A) of FIG. 8. Moreover, as in the case of the storage control unit 3 shown in FIG. 1, as indicated by (B) and (C) in FIG. 8, as the written-in address on the CM-A cache memory 31 and the written-in address on the CM-B cache memory 31, the following addresses (a1) and (a2) are respectively produced from the addressing information indicated by (A) in FIG. 8. That is,

(a1) (PA-A of CM-A) + (offset value); and

(a2) (PA-B of CM-B) + (offset value).

Thus, also in the storage apparatus 1 shown in FIG. 6, the data to be written is stored at different addresses in the master CM and mirror CM cache memories 31 and, in this case, the offsets (writing positions) are managed to be equal to each other in the respective cache pages regardless of the difference in page address.

The buffer 46 is located on the host interface module 20 side of the PCI bridge module 40, and temporarily retains

the data to be written received from the host interface module 20. The buffer 47 is provided in corresponding relation to each of the management modules 30, and receives the written data, to be transferred to the management module 5 30, from the target write buffer 46 to retain it temporarily.

The internal transfer control devices 45 and the transfer instruction devices 48 fulfill a function similar to that of the aforesaid data transfer control means 42, and in fact, these devices 45 and 48 are arranged in a state 10 integrated with each other. Each of the internal transfer control devices 45 transfers, on the basis of an output (two bus numbers in the aforesaid (5) and (6) fields) of the address analysis/conversion device 44, the data to be written from the host interface module side buffer 46 to 15 the management module side buffer 47, and the transfer instruction device 48 makes an instruction, on the basis of an output (the aforesaid addresses (a1), (a2)) of the address analysis/conversion device 44, to the management module side bus control device 49 for transmitting the 20 written data of the management module side buffer 47 to the designated address.

With the above-described arrangement, in the storage control unit 3 shown in FIGs. 6 and 7, when the host interface module 20 receives a data writing request as the I/O request 25 from the host 4 through the fiber channel interface bus 50, the address designation means 21 produces addressing information (see (A) in FIG. 8) for designating two

written-in destinations. This addressing information, together with the data to be written, is transferred from the host interface module 20 through the PCI bus 52 to the PCI bridge module 40 (see an arrow A21 in FIG. 7). At this  
5 time, the data to be written is transferred from the host interface module 20 in a dual write PCI address format.

In the PCI bridge module 40, when receiving the dual write, the same data to be written is simultaneously written through the buffer control device 43 in the two buffers  
10 46 with respect to one PCI writing (see arrows A22 in FIG. 7). Moreover, in the PCI bridge module 40, when recognizing a writing command including the addressing information, the address analysis/conversion device 44 extracts/produces two transferred-to addresses (PCI bus  
15 numbers for specifying the PCI buses of CM-A and CM-B) from the addressing information coming from the host interface module 20, and further produces written-in addresses (a1) and (a2) in each of the cache memories 31.

In addition, the internal transfer control devices  
20 45 transfer the written data retained in the two buffers 46 to the two buffers 47 respectively corresponding to the CM-A and CM-B PCI bus numbers (see arrows A23 in FIG. 7), and the transfer instruction devices 48 and the bus control devices 49 transfer the written data, transferred to the  
25 two buffers 47, to the CM-A and CM-B so that it is simultaneously written at the designated addresses (a1) and (a2) in the cache memories 31 (see arrows A24 in FIG.

7).

Thus, also in the storage control unit 3 shown in FIGs. 6 and 7, the forking writing (dual-writing of the data to be written) is carried out as indicated by the arrows A21 to A24 in FIG.7, and the same written data is written in a CBE of a cache page in the master area of the CM-A cache memory 31 and in a CBE (paired CBE) of a cache page in the mirror area of the CM-B cache memory 31 which is associated with the aforesaid master area cache page CBE in advance.

As described above, with the storage control unit 3 (storage apparatus 1) according to the embodiment of the present invention, when the host interface module 20 carries out the address designation once (single address designation) to perform once the data transfer (single data transfer) from the host interface module 20 to the PCI bridge module 40, the data to be written is transferred to the cache memories 31 of the two management modules 30 through the PCI bridge module 40, which enables a higher transfer rate than a standardized bus (PCI bus), to be doubly written therein. Therefore, the dual writing (forking writing) into the cache memories 31 can be made at a high rate, which achieves considerable improvement of the processing performance.

In addition, since the address designation means 21 of the host interface module 20 designates, as the written-in addresses for written data in the cache memories 31, a page address in the cache memory of each of the management modules

30 and an offset address in each page, the data to be written can be transferred to different addresses on the cache memories of the two management modules 30 to be written therein, which increases the degree of freedom of the cache management and achieves the efficient use of the cache memories 31.

Still additionally, since the management of the mirror cache can be made from the master CM side through the use of the management means (cache management firmware) 32 of each of the management modules 30, the address designation means 21 of the host interface module 20 can produce the addressing information on the basis of the information acquired from the management means 32 in one of the two management modules 30 without making communications between the two management modules 30 which are in mirror relation to each other.

Yet additionally, since the mirror cache can be managed from the master CM side through the use of the management means (cache management firmware) 32 of each of the management modules 30, one management module 30 (CM-A) can use the mirror area of the other management module 30 (CM-B) efficiently without making communications between the two management modules 30 which are in mirror relation to each other. For example, in a case in which the capacity of the master area of the cache memory 31 in one management module 30 (CM-A) runs short, a free area of the mirror area in the cache memory 31 of the other management module 30 (CM-B)



can be utilized on the basis of a management situation by the management means 32, which enables the efficient (approximately 100%) use of both the master area and the mirror area in the cache memory 31.

5           Moreover, since standardized buses (PCI buses 51 to 53) are used as the interfaces outside the PCI bridge module 40, the existing hardware can directly be used as the host interface module 20 and the management modules 30, and the forking writing (data dual writing) according to the present  
10 invention is realizable only by altering the address designation method (the function of the address designation means 21).

          It should be understood that the present invention is not limited to the above-described embodiments, and that  
15 it is intended to cover all changes and modifications of the embodiments of the invention herein which do not constitute departures from the spirit and scope of the invention.

          For example, although in the above-described  
20 embodiment the present invention is applied to a storage control unit, the present invention is not limited to this but, as in the case of the above-described embodiment, it is also applicable to an apparatus having a dual-writing function (multi-writing function) whereby data fed from  
25 an external apparatus such as a host is written two or more cache memories, and this can provide the effects similar to those of the above-described embodiment.

Moreover, although in the above-described embodiment the interface provided between the host 4 and the storage control unit 3 is a fiber channel interface, the present invention is not limited to this.

5           Still moreover, the functions of the above-mentioned address designation means 21, management means 32, address production means 41, data transfer control means 42, bus control device 43, address analysis/conversion device 44, internal transfer control device 45, transfer instruction  
10       device 48 and bus control device 49 can be realized in a manner such that a CPU (Central Processing Unit) which acts as the host interface module 20, the management modules 30 and PCI bridge module 40 executes a predetermined program. For example, this program can be offered in a state recorded  
15       in a computer-readable recording medium such as flexible disk, CD-ROM, CD-R, CD-RW or DVD.

As described above, according to the present invention, when the data transfer from the host interface module to the bridge module is made only once, the data to be written  
20       is transferred through the bridge module, which provides a higher transfer rate than a standardized bus, to two cache memories to be doubly written therein. Therefore, the dual-writing in the cache memories can be performed at a higher speed, thereby enhancing the processing performance  
25       considerably.

Accordingly, the present invention is suitable for use in an apparatus having a data dual-writing function,

such as a storage control unit made to control the access to a physical device (magnetic disk unit or the like) of a host, and the usability thereof is extremely high.